

každého učícího se jedince specifickou strukturu. Ta je určována zvláštnostmi žákových poznávacích procesů, charakterem jeho motivačních procesů, specifickou konkrétního sociálního prostředí a v neposlední řadě rovněž úrovní řízení učení (učitelem, technickými prostředky, kvalitou učebnic a jiného studijního materiálu).

K diagnostice studijního stylu se na středních a vysokých školách využívá dnes již známého dotazníku studijního stylu (původní verzi autorů N. Entwistla a P. Ramsdena přeložil a upravil J. Mareš). Tento materiál je auto-diagnostický, každý student si jej vyplní a vyhodnotí podle přiloženého klíče. Dotazník studijního stylu má pro studenta význam především na počátku studia. Za předpokladu, že je vyplněn pravdivě a sebekriticky, může přispět k odhalení zdrojů či příčin případných studijních obtíží a problémů.

V poslední době se u nás prováděly výzkumy učebních stylů žáků základních a středních škol. Byla vytvořena česká verze dotazníku LSI (Learning Style Inventory) autorů R. Dunnové a K. Dunna (Univ. St. John v New Yorku) a G. Price (Lawrence, Kansas). Původní materiál obsahuje kolem dvaceti proměnných a je určen pro žáky 3.–4. ročníku a 5.–12. ročníku. Na české verzi nezávisle na sobě pracovali I. Pýchová, H. Kantorková a J. Mareš. Ze tří překladů vznikla pracovní verze dotazníku ověřovaná u žáků českých základních škol. Česká verze odpovídá originálnímu materiálu a má dobrou reliabilitu (blíže Mareš a Skalská, 1994, s. 248n.).

V této kapitole jsme nabídli čtenářům některé základní diagnostické prostředky, jimiž lze poznat žáka z hlediska jeho zájmů, vztahu k předmětu a jeho hodnotových orientací. Pro učitele může být užitečné poznat úroveň žákova sebepojetí i jeho perspektivní orientaci. Především, že výše uvedené pohledy na žáka a třídu zdaleka nejsou dostačující. Mají učitelé poskytnout pouze inspiraci pro hlubší studium týkající se diagnostické problematiky i pro jejich vlastní nápady a tvořivost.

Je zřejmé, že není nutné a ani možné, aby každý učitel v takovém rozsahu diagnostikoval každého žáka. Je ale nutné, aby třídní učitel hlouběji poznal zejména žáky, kteří mají potíže. Měl by totiž umět poradit jim, jejich rodičům i učitelům. To bez důkladnějšího poznání žákovy osobnosti nelze.

7.2 Hodnocení vzdělávacích výsledků žáků

Miroslav Chráska

V našem stručném pojednání o hodnocení (evaluaci) vzdělávacích výsledků budeme věnovat pozornost zejména těm činnostem, které provádějí učitelé při své každodenní práci se žáky; nebudeme si tedy všimnat např. hodnocení práce samotných učitelů, evaluace vzdělávacích výsledků škol apod. K použití zde popsaných postupů ve výuce se vrátíme ještě v následující části knihy.

2.1 Jaký je současný stav hodnocení vzdělávacích výsledků žáků?

Hodnocení vzdělávacích výsledků žáků může učitel použít nejrůznějších metod. Současný systém hodnocení i klasifikace žáků je však založen převážně na *dividuálním ústním zkoušení*. Dobře připravená a provedená ústní zkouška má pro výuku velký význam zejména proto, že při ní vzniká jen těžko zastupitelný osobní kontakt mezi učitelem a žákem. Zkoušející může bezprostředně reagovat na nesprávné nebo nepřesné odpovědi, má možnost přesvědčit se, šlo o náhodnou chybu (např. o přeřeknutí), či o zásadní neznalost. Ústní zkouška má velký význam i pro rozvíjení vyjadřování a myšlení žáka. Ústní zkouška může být také komplexnější, protože neznalost jedné části učiva je možno za pomoci učitele překlenout a ve zkoušce pokračovat. Přes tyto neobjektivní klady má ústní zkouška v současné době řadu nedostatků.

Jaké nedostatky může mít ústní zkouška?

Velkým nedostatkem ústních zkoušek bývá to, že mívají mnohdy velmi **chudou myšlenkovou strukturu** (Fenclová, 1984). Taková zkouška nedává žákovi dostatek příležitostí, aby ukázal, zda pochopil základní vztahy v učivu a jak je schopen provádět konkrétní nebo abstraktní myšlenkové operace. Velmi často jsou ústní zkoušky zaměřují pouze na ověření zapamatování probíraného učiva. *Prádliční* ústní zkouška postihuje hlavně kvantum vědomostí, méně již jejich kvalitu. Vede spíše k memorování než k řešení problémů a k rozvoji myšlení. Často také obtížnostní struktura zkoušky bývá taková, že se v ní nemohou uplatnit žáci slabší i žáci velmi dobří.

Závažným nedostatkem ústního zkoušení je jeho **velká časová náročnost**. Zkoušení jednoho žáka se věnuje často až 10–15 minut. Tento čas většinou sami učitelé hodnotí jako ztrátový, protože pro výuku mnoho nepřináší.

Zkoušení je často chápáno učiteli jako „**nutné zlo**“, jen jako nezbytný podklad pro klasifikaci. Zkouška není chápána jako zpětná vazba mezi učitelem a žákem, ale odtrženě od procesu vyučování.

Zkoušky také **nemívají dostatečně jasný cíl**. Učitel si často nestanoví, co chce zkouškou zjistit. Mnohdy např. zkouší jen formální vědomosti, ale výsledku zkoušení přitom připisuje daleko širší platnost. To, že ústní zkouška nemá dostatečně jasný cíl, se velmi často projevuje tím, že jednotlivé úlohy zkoušky se zaměřují na celou řadu vzájemně nesouvisejících skutečností a ve svém celku neumožňují posouzení kvality vědomostí žáka.

Malá objektivita hodnocení je zdrojem většiny konfliktních situací mezi učitelem na jedné straně a žákem, rodiči, školskou správou na straně druhé. Slabá objektivita, která je jedním z nejzávažnějších nedostatků ústního zkoušení, je zřejmě způsobena řadou příčin. Na předním místě je to skutečnost, že hodnocení nevyjadřuje nikdy jen výkon žáka při zkoušce, ale je posudkem schopností a celkového dojmu, kterým žák působí (psychologové hovoří v této

souvislosti o tzv. preferenčních postojích). Při posuzování výsledku zkoušky se učitel snaží dosáhnout jednoty mezi hodnocením, které si přeje použít, a mezi skutečným výkonem žáka. Při vytváření preferenčních postojů u učitelů se uplatňuje řada vlivů, např. zevnějšek žáka, jeho chování, prospěch v ostatních předmětech, povahové vlastnosti žáka, ale třeba i sociální postavení rodičů, protekce atd.

Jinou, závažnou, příčinou nízké objektivity při ústním zkoušení je, že učitelé nemají pevné normy pro hodnocení. Hodnocení a klasifikace získané u jednoho učitele neodpovídají stejnému hodnocení a klasifikaci u jiného učitele. Hodnocení a klasifikace tedy nemají absolutní platnost. Ta se jim však bohužel často připisuje při přijímání žáků na vyšší typ školy, při přijímání do určitého povolání apod. Klasifikační řád sice rámcově vymezuje, jaké vědomosti a dovednosti by žák měl mít k dosažení jednotlivých klasifikačních stupňů, ale praxe je taková, že učitelé mají měřítko hodnocení velmi rozdílná.

Při ústním zkoušení se doporučuje (zejména začínajícím učitelům) postupovat podle určitého, předem připraveného scénáře. Písemný scénář zkoušky by měl obsahovat zejména přesné formulace otázek, příp. i předpokládané odpovědi žáků. Vedle věcné správnosti otázek a jejich přesnosti a srozumitelnosti by učitel měl zejména zvážit cíl, jehož má být dosaženo. Je třeba se zamyslet nejen nad obsahem zkoušeného tématu, ale také nad požadovanou úrovní osvojení poznatků. V tomto směru nám dobře poslouží znalost některé z taxonomií výukových cílů (např. Bloomova taxonomie, viz kapitola 10). Je žádoucí, aby se při zkoušení nezjišťovalo jen pouhé zapamatování poznatků, ale zejména vyšší cílové kategorie jako pochopení, aplikace, analýza, syntéza, příp. hodnocení. Při přípravě scénáře ústní zkoušky dbáme i na to, aby zkouška měla vhodnou obtížnostní strukturu, tj. aby obsahovala otázky různě obtížné, seřazené pokud možno od nejsnadnější po nejobtížnější. Důležitým hlediskem při přípravě ústní zkoušky je také časová úspornost. Na individuální ústní zkoušku jednoho žáka bychom měli plánovat maximálně deset minut času. Součástí přípravy ústní zkoušky by také mělo být rozhodnutí, které správné, resp. které nesprávné odpovědi předpokládáme pro dosažení jednotlivých klasifikačních stupňů. Dobře promyšlený scénář individuální ústní zkoušky představuje následující ukázka (Fuka a kol., 1985):

Ústní zkouška z fyziky

Téma: mechanická práce

1. Kdy konáme mechanickou práci?
2. Uveď dva praktické příklady.
3. Jaký účinek musí mít síla na těleso, aby se konala práce?

pokračování

pokračování

4. Kde se vzala síla, která uvádí těleso do pohybu?
5. Jaká je jednotka síly a dráhy?
6. Uveď vzorec pro výpočet práce a jednotku práce.
7. Co platí o směru síly při konání práce?
8. Uplatní se celá síla na práci, nesouhlasí-li směr síly se směrem pohybu?
9. Jak se změní velikost práce, působíme-li pětkrát menší silou po pětkrát větší dráze?
10. Jak vypočítáme sílu, známe-li vykonanou práci a dráhu?
11. $A = 16 \text{ J}$, $s = 4 \text{ m}$, $F = ?$ (z paměti).
12. Uveď dva příklady, při nichž se vykoná práce velikosti 30 J .
13. Lze na nějakém jednoduchém stroji ušetřit práci? Odůvodni.
14. Proč užíváme jednoduché stroje, když na nich práci neušetříme?

Všechny otázky v ukázce se týkají jednoho fyzikálního pojmu (mechanická práce), přičemž otázky 1, 5, 6 zkouší prostě zapamatování poznatků, zatímco ostatní se zaměřují na vyšší cílové kategorie.

Učitelům, kteří mají zájem na tom, aby jejich zkoušení probíhala adekvátně, lze doporučit pořízení záznamu svého zkoušení pomocí magnetofonu se skrytým mikrofonom. Záznam ústní zkoušky lze potom hodnotit např. pomocí následujících kritérií (Fenclová, 1984):

1. Co bylo tématem a cílem zkoušky?
2. Měla zkouška nějakou myšlenkovou strukturu? Kolik otázek zkoušky zjišťovalo jen zapamatování informace?
3. Jak byly formulovány úlohy zkoušky? Byly dostatečně srozumitelné, přesné, jednoznačné a pro žáky dosažitelné?
4. Byla zkouška pro žáky zajímavá, jak moc byla únavná?
5. Měla zkouška obtížnostní gradaci?
6. Dodávala zkouška slabším žákům odvalu, nebo je odrazovala? Zvyšovala jejich sebedůvěru? Mohla spíše zlepšit, nebo spíše zhoršit vztah žáků k předmětu?
7. Kolik času zkouška potřebovala a kolik informací poskytla učitelé?

Uvedených kritérií samozřejmě můžeme využít nejen při hodnocení vlastního zkoušení, ale i při hodnocení ústního zkoušení v rámci náslechnů a cvičných hospitací.

Tradiční ústní zkouška již nemůže dnes sama o sobě pro všestranné a vyvážené hodnocení žáka stačit. Velmi vhodným a ekonomickým doplňkem ústní zkoušky může být např. kvalitní didaktický test.

7.2.2 Co je to didaktický test?

Pojem didaktický test je sice u různých autorů definován odlišně, ale různá vymezení se shodují v tom, že se jedná o zkoušku, která se orientuje na objektivní zjišťování úrovně zvládnutí učiva u určité skupiny osob. Od běžné zkoušky se didaktický test ovšem liší zejména tím, že je navrhován, ověřován, hodnocen a interpretován podle určitých, předem stanovených pravidel. Stručná a výstižná je definice didaktického testu, jak ji uvádí P. Byčkovský (1982): Didaktický test je **nástroj systematického zjišťování (měření) výsledků výuky.**

Učitelé v praxi za test (resp. didaktický test) považují někdy jen krátkou písemnou zkoušku, někdy také jen zkoušku, která je sestavena výhradně z úloh s výběrem odpovědí. To je však zúžené chápání pojmu test (resp. didaktický test). V zahraničí se používají např. testy, jejichž vypracování žákovi nebo studentovi trvá i několik hodin (např. testy v závěru vysokoškolského studia některých oborů). Didaktický test také nemusí být nutně zkouškou písemnou. Existují a používají se testy řízení motorových vozidel, testy psaní na stroji atd. Konečně, didaktický test zdaleka ne vždy obsahuje jen úlohy s výběrem odpovědí. Mnohé testy užívají úloh, u nichž se požaduje řešení určitého problému, pojednání (esej) na dané téma apod.

7.2.3 Jaké jsou druhy didaktických testů?

V pedagogické praxi se můžeme setkat s didaktickými testy různé kvality a různého druhu. Jednotlivé druhy didaktických testů mají své specifické vlastnosti a liší se tím, jaké informace pomocí nich získáváme. Při popisu jednotlivých typů didaktických testů budeme postupovat podle klasifikace, kterou navrhl P. Byčkovský (1982).

Testy rychlosti

U těchto testů se zjišťuje, jakou rychlostí je žák schopen řešit určitý typ testových úloh. Testy rychlosti mají pevně stanovený časový limit pro řešení a obsahují velmi snadné úlohy. Předpokládá se, že všichni zkoušení žáci tyto úlohy zvládají a že se liší pouze v rychlosti řešení. Příkladem testu rychlosti je test rychlosti čtení, ve kterém měříme, kolik slov za minutu je žák schopen správně přečíst (aniž při tom přihlížíme ke kvalitě čtení), nebo test přepisu textu na psacím stroji, ve kterém měříme počet správných úhozů za minutu, apod.

Testy úrovně

Většina testů používaných v současné době na našich školách jsou testy, které se svým charakterem blíží testům úrovně. Čisté testy úrovně nepoužívají žádné časové omezení (časový limit) a výkon v nich je dán pouze úrovní vědomostí

nebo dovedností zkoušeného. Z praktických důvodů však bývá nutné s určitým velmi volným limitem pracovat zpravidla vždy.

Pokud testy úrovně používají časového limitu, pak je volen tak, aby znamenal přerušení práce jen pro ty nejpomalejší žáky. Výzkumy ukazují, že tito nejpomalejší žáci mají ve většině případů také nejslabší vědomosti a ani při dalším prodloužení času nedosahují lepších výsledků. Úlohy jsou totiž v testu zpravidla řazeny se vzrůstající obtížností, takže velmi pomalý žák v okamžiku přerušení práce na testu řeší ty nejobtížnější úlohy, které by byl již sotva schopen vyřešit.

Někdy i testy úrovně používají rychlosti jako vedlejšího kritéria pro hodnocení výkonu v testu. Je např. možno použít systému kombinovaného hodnocení, při kterém žákovi, který vyřeší správně většinu testových úloh (např. 80 % a více), připíšeme za každou „ušetřenou“ minutu jeden bod navíc.

Testy standardizované

Didaktické testy, které jsou připravovány důkladněji a které také mají úplnější vybavení, se označují jako testy standardizované. Standardizovaný didaktický test je připravován profesionálně, je důkladně ověřen, takže jsou známy jeho základní vlastnosti. Tyto testy vydávají většinou specializované instituce (u nás donedávna Psychodiagnostika Bratislava). Součástí příslušenství standardizovaného didaktického testu je testová příručka (manuál), ze které se uživatel dozví o vlastnostech testu, o jeho správném použití atd. Většinou je také k dispozici standard (testová norma) pro hodnocení dosažených výkonů.

Nestandardizované didaktické testy

Didaktické testy, u nichž nebyly realizovány všechny kroky obvyklé při přípravě a ověřování testů standardizovaných, označujeme jako testy nestandardizované (učitelské, neformální). Neproběhlo u nich ověřování na větším vzorku žáků, a nejsou tudíž známy všechny jejich vlastnosti. Tyto testy si připravují učitelé sami pro svoji vlastní potřebu. U testů nestandardizovaných není také k dispozici testová příručka ani objektivně stanovený testový standard (testová norma). I při konstrukci těchto testů by však učitelé měli dbát pravidel a zásad, které se doporučují u testů standardizovaných.

Někdy se užívá i termínu *testy kvazistandardizované*, čímž se rozumí testy připravované dokonaleji než testy učitelské, u nichž ale standardizace nebyla provedena bezezbytku. Kvazistandardizovaným testem je např. didaktický test, zjišťující úroveň vědomostí žáků v daném předmětu na určité škole (několik paralelních tříd) nebo na několika školách. Konstrukci těchto testů bývá většinou věnována větší pozornost než u testů nestandardizovaných, bývají známy některé jejich vlastnosti a někdy bývají k dispozici i standardy pro hodnocení testových výsledků.

Testy kognitivní a testy psychomotorické

Dělení didaktických testů na kognitivní a psychomotorické vychází z dělení lidského učení do tří oblastí podle B.S. Blooma (učení kognitivní, afektivní a psychomotorické). Výsledky učení afektivního se didaktickými testy nezjišťují (k tomuto účelu se používají dotazníky, různé škály apod.). Pokud didaktický test měří úroveň (kvalitu) poznání u žáků, jedná se o test kognitivní, pokud testem zjišťujeme výsledky psychomotorického učení, jde o test psychomotorický. Příkladem kognitivních testů jsou testy, ve kterých má žák řešit úlohy z matematiky, překládat text do cizího jazyka atd. Příkladem testu psychomotorického je test psaní na stroji. V současné praxi se daleko častěji používají testy kognitivní, užití testu psychomotorického je spíše výjimkou.

Testy výsledků výuky a testy studijních předpokladů

V běžné pedagogické praxi se doposud téměř výlučně používají didaktické *testy výsledků výuky*, které měří to, co se žáci v dané oblasti naučili. *Testy studijních předpokladů* měří úroveň obecnějších charakteristik jedince, které jsou potřebné k dalšímu studiu. Testy studijních předpokladů by se měly používat zejména při přijímání žáků ke studiu na vyšší typ školy. Zatím však je praxe taková, že testy, které se k tomu účelu na našich školách používají, se příliš neliší od běžných testů, tj. od testů výsledků výuky. Konstrukce testů studijních předpokladů je podstatně náročnější a vyžaduje vedle pedagogické kvalifikace autora také dobrou kvalifikaci psychologickou.

Testy rozlišující (testy relativního výkonu)

Podle toho, jakým způsobem interpretujeme (vysvětlujeme a hodnotíme) výkon žáka v testu, můžeme rozlišit tzv. didaktické testy rozlišující (testy relativního výkonu) a didaktické testy ověřující (testy absolutního výkonu). Testy rozlišující se také označují jako testy statisticko-normativní nebo jako NR testy (norm-referenced tests). Hlavní rozdíl mezi těmito dvěma druhy testů spočívá v tom, že u testů rozlišujících se výkon žáka určuje vzhledem k populaci testovaných, zatímco u testů ověřujících se výkon určuje vzhledem ke všem možným úlohám, které určité učivo reprezentují.

V naší pedagogické praxi se zatím používají téměř výlučně testy rozlišující. Základní ideou, o kterou se opírá koncepce těchto testů, je snaha dosáhnout maximální možné objektivity a diferencovanosti hodnocení testových výkonů. Výkon žáka v testu je srovnáván s výkony ostatních žáků, v případě standardizovaných rozlišujících testů s výkony celé žákovské populace. Rozlišující didaktické testy jsou tedy konstruovány tak, že umožňují rozhodnout, jaký výkon v testu žák dosáhl vzhledem k celé populaci, k níž patří. Umožňují posoudit, zda určitý konkrétní žák je ve srovnání s ostatními žáky „velmi slabý“, „podprůměrný“, „průměrný“ atd.

Testy ověřující (testy absolutního výkonu)

Didaktické testy ověřující jsou v literatuře označovány také jako kritériální testy nebo CR testy (criterion-referenced tests). Úkolem ověřujících testů je prověřit úroveň vědomostí a dovedností žáka v přesně vymezené oblasti (části učiva). Výkon testovaného se přitom nesrovnává s výkonem jiných žáků (populace), nýbrž se vyjadřuje vůči všem úlohám, které reprezentují dané učivo.

U ověřujících testů je kritériem úspěchu předem stanovený stupeň zvládnutí učiva. Ověřující testy požadují u vybraných základních poznatků téměř úplné zvládnutí (četnost chyb musí zhruba odpovídat náhodě). Ověřující testy nesilují o diferencované hodnocení žáků, jak je tomu u testů rozlišujících. Cílem je v podstatě rozhodnout, zda žák zvládl učivo, nebo nikoli. Při konstrukci ověřujících testů je základním problémem výběr učiva, které musí žák bezpečně zvládnout. Toto učivo se potom transformuje do testových úloh. Aby se bezpečně ověřilo zvládnutí určitého učiva, požaduje se, aby každý testovaný žák byl pokryt větším počtem testových úloh.

V naší pedagogické praxi se ověřující testy zatím téměř neuvžívají. Jejich teorie se však rychle rozvíjí a zdá se, že rozšíření tohoto druhu testů bude pro testování vědomostí žáků přínosem.

Testy vstupní, průběžné a výstupní

Vstupní didaktické testy se zadávají na začátku výuky určitého celku učební látky a jejich cílem je postihnout úroveň vědomostí a dovedností, které jsou pro úspěšné zvládnutí daného celku učiva důležité. Zařazení vstupního testu na začátku výuky může posloužit jako zdroj cenných informací např. v případě, že hodláme realizovat diferencovanou výuku.

Průběžné didaktické testy se zadávají v průběhu výuky a jejich posláním je poskytovat učiteli zpětnovazební informace potřebné k optimálnímu řízení výuky. Obvykle zkouší jen poměrně malou část učiva a jejich posláním je sledovat, jak žáci probírané učivo přijímají, chápou a jak si je osvojují. V této souvislosti se často hovoří o tzv. *formativních testech*, které slouží ke sledování procesu formování vědomostí a dovedností u žáků. Tyto testy neslouží většinou k hodnocení žáků, nýbrž k hodnocení průběhu výuky.

Výstupní didaktické testy se zadávají na konci výukového období nebo na konci určitého celku a většinou poskytují informace potřebné pro hodnocení žáků. Bývají také označovány jako *testy sumativní*.

Testy monotematické a polytematické

Monotematické testy zkoušejí jediné téma učební látky, testy polytematické zkoušejí učivo několika tematických celků. Testy polytematické jsou proto náročnější z hlediska přípravy i konstrukce.

Testy objektivně skórovatelné

Testy objektivně skórovatelné obsahují úlohy, u nichž lze objektivně rozhodnout, zda byly řešeny správně, či nikoli. Výhodou těchto testů je, že skórování může provádět jakákoli osoba (někdy i stroj).

Vzhledem k tomu, že velká většina používaných didaktických testů se vyznačuje možností objektivního skórování, vznikla u části pedagogické veřejnosti nesprávná představa, že test je zkouška, která vždy obsahuje pouze objektivně hodnotitelné úlohy (např. úlohy, kde žák vybírá správnou odpověď nebo úlohy, kde žák formuluje vlastní, ale velmi stručnou, a tudíž objektivně hodnotitelnou odpověď).

Testy subjektivně skórovatelné

Subjektivně skórovatelné testy (označované často jako esej testy) obsahují úlohy, u nichž není možno stanovit jednoznačná pravidla pro skórování. Mezi subjektivně skórovatelné testové úlohy patří např. tzv. otevřené široké úlohy, ve kterých žák volně odpovídá na položenou otázku uvedením rozsáhlejší odpovědi. Ukazuje se, že není rozumné vyhýbat se používání takových úloh jen proto, že neumožňují objektivní skórování. Otevřené široké úlohy totiž mohou zkoušet daleko komplexnější vědomosti a dovednosti než úlohy objektivně skórovatelné.

7.2.4 Jak vytvořit dobrý didaktický test?

Konstrukci didaktického testu bychom neměli začínat přímo navrhováním testových úloh. Tento postup nejčastěji vede k tomu, že vytvoříme testové úlohy, které se sice snadno navrhnou, ale ve svém celku vytvářejí nevyvážený didaktický test, který nepokrývá rovnoměrně celé učivo a zaměřuje se většinou na pouhou reprodukci zapamatovaných poznatků. Nejsnadněji se totiž navrhnou takové úlohy, které zkoušejí pouhé zapamatování látky. Těchto úloh bývá u autorů nedostatečně poučených velká převaha.

Plánování didaktického testu

První otázkou, kterou by si měl autor testu položit, je: K jakému účelu má didaktický test sloužit? Účelem testu může být např. zjištění výsledků výuky na konci tematického celku nebo na konci pololetí či roku, zjištění, jak žáci probírané učivo přijímají a chápou. Didaktický test může také sloužit ke kontrolním nebo inspekčním účelům, k výběru vhodných studentů na vyšší typ školy atd.

Po ujasnění účelu testování se zpravidla *rámcově vymezuje obsah* testu.

Pohyb a síla. Vzájemný pohyb těles. (Učivo 31 hodin fyziky, 7. ročník základní školy)

Skladba. Tvarosloví. (Učivo 40 hodin českého jazyka, 6. ročník základní školy)

Rámcově vymezený obsah testu je třeba upřesnit (*specifikovat*) tak, aby bylo patrné, jaký obsah mají jednotlivé úlohy zkoušet. Zejména je třeba určit:

- jakou úroveň osvojení vědomostí mají jednotlivé úlohy zkoušet,
- kolik úloh má zkoušet jednotlivé prvky učiva,
- kolik úloh musí obsahovat celý test.

V literatuře (např. Byčkovský, 1982) lze nalézt několik technik, kterými lze tuto specifikaci provést. Na tomto místě se omezíme jen na uvedení hlavních zásad, jimiž je specifikace určována.

Je žádoucí, aby jednotlivé testové úlohy nezkoušely jen pouhé zapamatování učiva, ale i vyšší cílové kategorie, např. porozumění poznatkům, používání poznatků v problémových situacích. V tomto směru jsou dobrou oporou pro autora testu různé taxonomie výukových cílů (srov. kapitola 10) pro kognitivní oblast. Pro plánování didaktických testů je pro svoji jednoduchost zvlášť vhodná taxonomie B. Niemiérka.

Při úvahách o celkovém počtu úloh v testu vycházíme z toho, že jedině test s dostatečným počtem úloh může poskytnout učiteli spolehlivou (reliabilní) informaci. V praxi lze považovat za spodní hranici počtu úloh v testu deset úloh. Testy s menším počtem úloh již nemohou (ani v případě sebedokonaleji vytvořeného testu) učiteli přinést dostatečně věrohodné informace. Horní hranice počtu úloh v testu je dána zejména časovými možnostmi. U testů monotematických bývá zpravidla maximálně kolem dvaceti úloh.

Jednotlivé prvky učiva by měly být v didaktickém testu reprezentovány takovým počtem úloh, který odpovídá jejich významu (důležitosti) a rozsahu.

Konstrukce didaktického testu

Konstrukcí didaktického testu rozumíme vytvoření jednotlivých testových úloh a vytvoření prvního návrhu (prototypu) didaktického testu.

Testovou úlohou rozumíme otázku, úkol nebo problém obsažený v testu. Kromě termínu testová úloha se často v literatuře užívá výrazu testová položka nebo testový úkol, v praxi často termínu otázka, úkol nebo příklad. Na kvalitě testových úloh závisí v podstatné míře kvalita celého testování. Navrhování testových úloh je velmi náročná činnost, k jejímuž úspěšnému provádění je třeba vedle zkušeností i náležitého teoretického poučení. Autor didaktického testu by měl být dobrým odborníkem v předmětu, pro který test připravuje, ale měl by být také dobrým pedagogem a psychologem, aby se dokázal vcítit do pozice žáků, které chce testovat.

Závažným rozhodnutím, před kterým autor testu v této fázi tvorby stojí, je, jaký typ úloh v didaktickém testu použít. Každý druh testových úloh má určité vlastnosti, výhody i nevýhody. O tom, který druh testových úloh použijeme, rozhoduje cíl, jež má testování plnit, obsah učiva, jež má být předmětem testování, materiální a technické podmínky, ale v neposlední řadě i obliba určitého druhu testových úloh u autora testu.

Nejdůležitější typy testových úloh uvedeme ve členění podle P. Byčkovského (1982).

Otevřené široké úlohy

Podle způsobu, jakým žák v testové úloze odpovídá, lze rozlišit úlohy otevřené (někdy v literatuře označované jako úlohy s tvořenou odpovědí nebo volnou odpovědí) a úlohy uzavřené (s nabízenou odpovědí, s nucenou volbou odpovědi). Otevřené úlohy lze dále rozdělit podle rozsahu požadované odpovědi na široké a se stručnou odpovědí.

V otevřených širokých úlohách se požaduje od žáka rozsáhlejší odpověď (např. půl strany nebo i delší). Může se např. požadovat pojednání na určité téma (např. *Význam díla K. H. Máchy pro českou poezii* nebo *Které byly hlavní příčiny vzniku první světové války?*), vyřešení určitého problému (např. *Navrhněte postup, kterým je možno určit hustotu neznámé kapaliny.*), popis určitého procesu (např. *Popište činnost čtyřdobého zážehového motoru.*). Požadovaný rozsah odpovědi se žákovi naznačuje velikostí vynechaného místa v testovém zadání.

Někdy bývá u otevřených širokých úloh vhodné vymezení strukturu požadované odpovědi:

Výroba surového železa (uvedte hlavní používané suroviny, nakreslete schéma vysoké pece a popište hlavní probíhající chemické reakce).

Někdy není nutno strukturu požadované odpovědi vymezovat, protože vyplývá z konvence, kterou by zkoušený měl znát, např.:

Popište hlavní stadia tělesného vývoje dítěte do šesti let věku.

Otevřené široké úlohy lze doporučit zejména při zkoušení komplexních vědomostí nebo dovedností, osvojovaných v delším časovém období. Jsou vhodné pro zkoušení vyšších úrovní osvojení učiva (např. řešení problémových situací), zatímco pro zkoušení nižších úrovní osvojení učiva (např. zapamatování) jsou výhodnější testy s úlohami objektivně skórovatelnými.

Široké testové úlohy se poměrně snadno navrhnou, ale jejich hlavní nevýhodou je nemožnost objektivního skórování. To však neznamená, že by se neměly používat. Komplexní vědomosti a dovednosti lze sice rozložit na dovednosti dílčí (které se dají zkoušet mnohem snadněji), avšak zvládnutí dílčích dovedností ještě nemusí znamenat zvládnutí dovedností komplexních.

Při skórování širokých úloh se často postupuje tak, že za správné a úplně zodpovězení úlohy se přisuzuje určitý počet bodů (např. 10). Za každou chybnou nebo chybějící část odpovědi se potom strhává určitý počet bodů. V některých

případech lze pro skórování otevřené testové úlohy vypracovat detailní předpis, který umožňuje téměř objektivní skórování.

Testy vytvořené z otevřených širokých úloh se často označují jako *esej testy*. Tyto testy obsahují zpravidla jen několik úloh. Od běžných písemných zkoušek se esej testy liší tím, že při jejich konstrukci, hodnocení i interpretaci výsledků se využívají všechna základní pravidla a postupy obvyklé u ostatních didaktických testů.

Úlohy se stručnou odpovědí

Úlohy se stručnou odpovědí požadují od žáka vytvoření a uvedení vlastní krátké odpovědi. Může se požadovat např. uvedení čísla, značky, symbolu, vzorce, jednoduchého grafu, určitého slova, příp. několika slov či krátké věty. Úlohy se stručnou odpovědí mohou být dvojího typu: *produkční a doplňovací*.

Příklady produkčních testových úloh:

Které jsou tři základní složky lidské potravy?

1.
2.
3.

Napište Archimedův zákon

.....

Příklady doplňovacích testových úloh:

Hlavním městem Švýcarska je

Po smrti českého krále Karla IV. v roce nastoupil na trůn jeho syn

Mezi výhody úloh se stručnou odpovědí lze počítat zejména to, že se snadno navrhnou. Další jejich výhodou je, že neumožňují žákům tak snadno uhodnout správnou odpověď bez příslušných vědomostí, jako je tomu u úloh s výběrem odpovědí. Většinou se předpokládá, že vytvoření odpovědi je pro žáka náročnější než pouhé rozpoznání správné odpovědi mezi nabídnutými alternativami.

Nevýhodou úloh se stručnou odpovědí je to, že žák mnohdy odpovídá správně, ale jinak, než si představoval autor testu. Opravování (skórování) testu, ve kterém jsou použity úlohy se stručnou odpovědí, proto nemůže provádět laik, nýbrž jen odborník, který zkoušenému učivu dokonale rozumí.

Úlohy dichotomické

U dichotomických testových úloh jsou žákovi předkládány dvě alternativy odpovědi s tím, že jedna je správná a tu má označit (např. podtržením, zakroužkováním). Tyto úlohy bývají také často označovány jako úlohy s dvoučlennou volbou nebo jako alternativní úlohy. Nejčastější formy dichotomických úloh uvádějí následující příklady:

Mistr Jan Hus byl upálen roku 1515.	ano – ne
Při vypařování kapaliny se teplo	spotřebovává – uvolňuje.

Výhodou dichotomických úloh je, že se velmi snadno navrhují. Tato konstrukční jednoduchost však může svádět k testování jednotlivých detailů, pouhých faktů (někdy i triviální povahy). Nedostatkem dichotomických úloh je velká pravděpodobnost uhodnutí správné odpovědi i bez příslušných vědomostí. Aby se věrohodnost výsledků získaných testem s dichotomickými úlohami zvýšila, je třeba, aby test obsahoval těchto úloh dostatečný počet.

Úlohy s výběrem odpovědí

Úlohy s výběrem odpovědí (v literatuře označované také jako úlohy s vícečlennou či vícenásobnou odpovědí, úlohy polynomické) vděčí za svoji teoretickou rozpracovanost především rozvoji programovaného učení (hlavně tzv. větvených programů). Úloha s výběrem odpovědí se skládá ze dvou částí: problému nebo otázky (tzv. kmenu úlohy) a nabídnutých odpovědí.

Úlohy s výběrem odpovědí se v testech vyskytují v několika formách.

a) Úlohy typu „jedna správná odpověď“

Základní formou je úloha, ve které žák vybírá *jednu správnou odpověď* z několika nabídnutých alternativ.

Cesta vlakem nám velmi rychle uběhla.

Podtržený větný člen je:

- A podmět
- B předmět
- C přívlastek
- D příslovečné určení

b) Úlohy typu „jedna nejpřesnější odpověď“

Jinou formou úloh s výběrem odpovědí jsou úlohy, kde se požaduje nejlepší nebo nejsprávnější odpověď. Takové úlohy mohou být pro žáky velmi obtížné, obtížnější než odpovídající úlohy otevřené.

Které z následujících tvrzení nejlépe odpovídá na otázku:
„Co je chemický prvek?“

- A Prvek je látka, která se skládá z atomů stejného druhu.
- B Prvek je látka, kterou již dále nelze dělit.
- C Prvek je látka složená z atomů, které mají stejné protonové číslo.
- D Žádné z předchozích tvrzení není naprosto správné.

c) Úlohy typu „jedna nesprávná odpověď“

V některých případech lze také požadovat uvedení *nesprávné odpovědi*. V tomto případě je ovšem nutné zápor ve kmenu úlohy patřičně zdůraznit, protože jinak může snadno dojít k přehlédnutí a žák odpoví špatně přesto, že má příslušné vědomosti.

Který z následujících dějů *není* formou oxidačního procesu?

- A dýchání
- B hnití
- C destilace
- D rezivění

d) Úlohy s vícenásobnou odpovědí

Jestliže má žák v testové úloze vybrat několik správných odpovědí, hovoříme o tzv. *vícenásobné odpovědi*. Jestliže se rozhodneme pro použití této úlohy, je nutné na to žáky předem upozornit. V úlohách, kde se požaduje výběr jen jedné odpovědi, je totiž výběr většího počtu odpovědí považován za chybu, a žáci by proto mohli váhat, zda více odpovědí uvést.

Kterými státy protéká (nebo kterých se alespoň dotýká) řeka Odra?

- A Německo
- B Rusko
- C Česká republika
- D Slovensko
- E Polsko

Určité problémy u úloh s vícenásobnou odpovědí nastávají při jejich skórování. Neexistuje zde totiž jen jedna naprosto správná a jedna naprosto nesprávná odpověď, ale také několik částečně správných odpovědí. Lze doporučit přístup, který lze stručně vyjádřit slovy „všechno, anebo nic“. Podle tohoto přístupu přidělíme 1 bod v případě, kdy žák označí všechny správné odpovědi, a 0 bodů tehdy, jestliže bude (třeba jen jedna) odpověď nesprávná.

e) **Situační úlohy**

Zvláštní modifikací testových úloh s výběrem odpovědí jsou úlohy označované někdy jako úlohy *situační* či *interpretační*. Jsou to úlohy, u nichž žák vybírá z podstatně většího počtu nabídek, než je obvyklé, přičemž nabídky nejsou předkládány ve formě dlouhého a nepřehledného seznamu, nýbrž vyplynou přímo z dané situace. Pravděpodobnost uhodnutí správné odpovědi bez příslušných vědomostí je u tohoto typu úloh zpravidla velmi malá.

Na místo označené hvězdičkou napište takovou číslici, aby výsledné šesticiferné číslo bylo dělitelné sedmi:

823*43

(V uvedené testové úloze nejsou na první pohled nabízeny žádné alternativy odpovědi, ve skutečnosti však žák v případě neznalosti učiva vybírá z deseti číslic. Pravděpodobnost náhodné volby správné odpovědi je v tomto případě 10 %.)

Častou námitkou proti užívání testových úloh s výběrem odpovědi bývá, že žák může mnohdy uhodnout správnou odpověď, aniž by učivo náležitě zvládl. Někteří pedagogové dokonce vidí v tomto způsobu testování jen jakousi *inteligentní hru*, při které žák s větší či menší dávkou štěstí hádá správné odpovědi. Podle našeho mínění takovéto názory zkreslují skutečnost. Kromě toho lze vhodně volenými opatřeními vliv „hádání“ omezit na minimální míru. (Podrobněji srov. Chráska, 1998.)

Přířazovací úlohy

Přířazovací úlohy obsahují instrukci a dvě množiny pojmů. Úkolem žáka je správně přiřadit pojmy jedné množiny k pojmům množiny druhé.

K názvům států v levém sloupci správně přiřaďte názvy jejich hlavních měst z pravého sloupce!

Švýcarsko ()	A Oslo
Norsko ()	B Dublin
Island ()	C Bern
Finsko ()	D Berlín
Nizozemsko ()	E Reykjavík
	F Helsinky
	G Amsterdam

V uvedeném příkladě je počet pojmů v pravém sloupci úmyslně větší než počet pojmů v levém sloupci. Tuto zásadu bychom měli u tohoto typu úloh dodržovat vždy. Pokud by totiž obě množiny pojmů byly stejně početné, měl by žák, který některá přiřazení zná, situaci usnadněnu tím, že by se mu počet možných přiřazení zmenšoval.

Výhodou přiřazovacích testových úloh je, že omezují možnost uhodnutí správné odpovědi na minimální míru. Jejich použití je však možné jen v poměrně omezeném okruhu učiva.

Uspořádací úlohy

V uspořádacích testových úlohách se od žáka požaduje, aby uspořádal prvky dané množiny pojmů jedné třídy do řady. Úloha tohoto typu se skládá z dané množiny prvků a z instrukce, která uvádí, podle kterého kritéria a jakým způsobem se mají prvky uspořádat. Prvky je možno řadit podle velikosti, významu, stupně obecnosti, chronologicky atd.

Seřad'te následující racionální čísla podle velikosti tak, že k nejmenšímu z nich připsáte 1 a k největšímu 4.

0,5	$\frac{12}{18}$	-0,001	$\frac{15}{60}$
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Určité problémy působí skórování uspořádacích úloh. Nesprávné seřazení prvků ve skupině může totiž být provedeno mnoha způsoby, přičemž se jedná o různé závažné chyby. Nejjednodušší způsob skórování se provádí tak, že za zcela správné vyřešení úlohy, tj. za uvedení naprosto správného pořadí, se přiděluje 1 bod, za všechna ostatní řešení 0 bodů.

Návrh prototypu didaktického testu

Zkušenosti ukazují, že bývá prospěšné vytvořené testové úlohy na nějaký čas (alespoň několik dnů) po návrhu odložit a teprve potom se k nim znovu kriticky vrátit. Z úloh, které při opakovaném posuzování (příp. při posuzování jinými odborníky) obstály, sestavíme první návrh (prototyp) didaktického testu.

Konstruujeme-li rozlišující test, potom je třeba, abychom úlohy v testu seřadili podle odhadované vzrůstající obtížnosti.

Součástí přípravy testu k jeho použití je také přibližné určení času, který budou žáci k vypracování potřebovat. U jednodušších otevřených úzkých úloh a u jednodušších úloh s výběrem odpovědi lze orientačně počítat s časem od 0,5 minuty do 1,5 minuty na jednu úlohu. V každém případě časový limit stanovujeme velmi volně a v běžných případech necháváme žáky pracovat prakticky bez časového omezení.

7.2.5 Jak lze ověřit kvalitu vytvořeného didaktického testu?

I když plánování a konstrukci didaktického testu věnujeme značnou pozornost (postupujeme poučeně a pečlivě), přesto si nemůžeme být nikdy jisti tím, jaké vlastnosti bude nakonec test mít. Relativně definitivní představu o vlastnostech testu můžeme získat až po důkladném vyzkoušení (ověření) testu na vzorku žáků. Toto ověřování se ovšem neprovádí pouze za účelem získání informací o kvalitě vytvořeného testu, nýbrž zejména proto, abychom mohli případně nevhodné vlastnosti testu odstranit, zmírnit nebo korigovat. Jestliže vytváříme didaktický test jen pro vlastní potřebu (nestandardizovaný test), potom zpravidla vystačíme s ověřením testu u těch žáků, které vyučujeme. U nestandardizovaných testů také autor většinou vystačí s pouhým odhadem vlastností testu a nemusí provádět jejich přesné stanovení. U testů standardizovaných se však každý uživatel s přesnými údaji o vlastnostech testu setkává, a je proto třeba, aby jim rozuměl a dokázal je správně interpretovat.

Ověřováním vytvořeného didaktického testu můžeme získat informace jednak o vlastnostech z jednotlivých úloh v testu, jednak o vlastnostech testu jako celku (Chráska, 1999).

Jaké vlastnosti by měla mít dobrá testová úloha?

Při posuzování vlastností jednotlivých testových úloh je věnována pozornost zejména tomu, jak dalece jsou vytvořené úlohy pro žáky *obtížné* a jakou mají tzv. *citlivost* (Mráz, 1977).

Obtížnost testových úloh

To, jak dalece je testová úloha pro žáky obtížná, je jednou z jejích základních charakteristik. Obtížnost jednotlivých testových úloh můžeme posoudit podle toho, kolik žáků je dokáže správně vyřešit. Při analýze obtížnosti se vypočítává buď *hodnota obtížnosti Q*, nebo *index obtížnosti P*.

Hodnota obtížnosti udává procento žáků ve vzorku, kteří danou úlohu zodpověděli nesprávně nebo vynechali:

$$Q = 100 \frac{n_n}{n},$$

kde Q je hodnota obtížnosti, n_n je počet žáků, kteří odpověděli v dané úloze nesprávně nebo neodpověděli vůbec, a n je celkový počet testovaných žáků.

Index obtížnosti je procento žáků ve skupině, kteří danou úlohu zodpověděli správně.

$$P = 100 \frac{n_s}{n},$$

kde P je index obtížnosti, n_s počet žáků, kteří odpověděli v dané úloze správně, a n je celkový počet žáků.

Za velmi obtížné lze pokládat testové úlohy, u nichž hodnota obtížnosti Q je vyšší než 80. Velmi snadné jsou naopak ty úlohy, které vykazují hodnotu obtížnosti Q nižší než 20. Velmi obtížných ani velmi snadných úloh by nemělo být v testu příliš mnoho. Úlohy extrémně obtížné, u nichž se hodnota obtížnosti Q blíží ke 100, jsou nevyhovující, a je nutno je z testu vyloučit. Úlohu extrémně snadnou, u níž se hodnota obtížnosti Q blíží k nule, je možno z psychologických důvodů doporučit jako úvodní úlohu v testu. Může totiž přispět k uklidnění žáků a k vytvoření potřebného pocitu jistoty. Zkušenosti ukazují, že nejvhodnější vlastnosti mají testové úlohy s hodnotou obtížnosti kolem $Q = 50$ (platí pro testy rozlišující).

Citlivost testových úloh

Citlivost úloh bývá často označována také jako *rozlišovací hodnota*, *diskriminační hodnota*, *rozlišovací ostrost* nebo jako *rozlišovací schopnost úloh*. Vysokou citlivost má taková úloha, kterou řeší s velkým úspěchem žáci, kteří mají celkově lepší vědomosti, zatímco žáci, kteří mají celkově horší vědomosti, v této úloze dosahují výsledků špatných. Citlivost úlohy tedy vyjadřuje, jak dalece daná úloha zvýhodňuje žáky, mající lepší vědomosti, před žáky, kteří mají vědomosti horší. K rozlišení žáků na žáky „s lepšími vědomostmi“ a na žáky „s horšími vědomostmi“ se většinou používá celkových výsledků ověřovaného didaktického testu.

Citlivost úlohy se dá exaktně posoudit pomocí výpočtu některého z koeficientů citlivosti, kterých byla navržena celá řada. Všechny tyto koeficienty mohou nabývat hodnot od -1 přes nulu do $+1$, přičemž platí, že čím vyšší hodnotu koeficient má, tím lépe úloha rozlišuje mezi žáky s lepšími vědomostmi a mezi žáky s horšími vědomostmi. Pokud koeficient citlivosti dosahuje hodnoty 0, znamená to, že úloha vůbec nerozlišuje mezi oběma skupinami žáků (žáci s lepšími i žáci s horšími vědomostmi jsou v této úloze stejně úspěšní). Záporné hodnoty koeficientu citlivosti vypovídají o tom, že úloha zvýhodňuje žáky, kteří mají v testu celkově horší výsledky. Kladné hodnoty koeficientu citlivosti naopak vypovídají o tom, že v úloze dosahují lepších výsledků žáci, kteří mají v testu lepší celkové výsledky (úloha zvýhodňuje žáky, kteří mají lepší vědomosti). Za přijatelné jsou většinou považovány hodnoty koeficientu citlivosti větší než $+0,20$. Podrobnější poučení o metodách výpočtů koeficientů citlivosti nalezne čtenář např. v práci Chráska (1998).

Velice nízkých (příp. záporných) hodnot nabývá koeficient citlivosti např. v úlohách, které jsou příliš komplikovaně formulovány a kdy obě skupiny žáků (skupina „lepších“ a skupina „horších“) mohou volit rozdílné strategie řešení. Zatímco žáci „lepší“ se snaží k řešení komplikované úlohy dojít složitými úvahami, při kterých se snadno dopouštějí chyb, žáci „horší“ se složitými úvahami příliš neznepokojují a odpověď s větším či menším štěstím hádají. Nízké hodnoty koeficientu citlivosti se objevují také u některých velmi obtížných úloh s výběrovými odpověďmi nebo u úloh, kde se zkouší formálně zvládnuté učivo.

Při posuzování citlivosti úloh u nestandardizovaných didaktických testů není třeba vypočítávat koeficienty citlivosti. Postačí, když si všimneme, jak v jednotlivých testových úlohách odpovídali žáci celkově „lepší“ a žáci celkově „horší“. Jestliže zjistíme, že v některé úloze dosáhli lepších výsledků spíše „horší“ žáci, znamená to, že s úlohou není něco v pořádku a že zřejmě má malou (nebo zápornou) citlivost.

Jaké vlastnosti by měl mít dobrý didaktický test jako celek?

Při posuzování testu jako celku se zpravidla věnuje pozornost jeho validitě a reliabilitě (Mráz, 1977).

Validita didaktického testu

Test je validní tehdy, pokud se jím zkouší skutečně to, co má být zkoušeno. U testů studijních výsledků zkoumáme, jak dalece se shoduje obsah testu s cílem a obsahem vyučování. V těchto případech nám jde především o tzv. obsahovou validitu testu. Obsah úloh didaktického testu by měl být v těchto případech reprezentativním vzorkem zkoušeného učiva. Pro testy výsledků výuky jsou proto kritériem validity příslušná kurikula vyučovacích předmětů.

Při posuzování validity u testů studijních předpokladů je situace podstatně složitější, protože je nutno zvážit, ze kterých vědomostí nebo dovedností je možno usuzovat na budoucí úspěšnost ve studiu. Testy studijních předpokladů by měly mít především tzv. predikční validitu, tj. schopnost předpovídat budoucí úspěšnost v učení.

Posouzení stupně validity testu se v praxi většinou ponechává na příslušném odborníkovi (ještě lépe na skupině odborníků). Jen výjimečně je možno posoudit validitu nově vytvořeného testu tak, že se srovnávají v něm dosažené výsledky s výsledky nějakého jiného testu, jehož validita je nesporná.

Reliabilita didaktického testu

Aby byl didaktický test reliabilní, je třeba, aby byl spolehlivý. Spolehlivost spočívá v tom, že za týchž podmínek by měl poskytovat stejné (velmi podobné) výsledky. Další podmínkou dobré reliability je přesnost testu. Didaktický test je přesný tehdy, jestliže při jeho použití nedochází k velkým chybám měření. Oba požadavky, tj. spolehlivost a přesnost, zahrnujeme pod společný pojem reliabilita.

K exaktnímu posouzení míry reliability didaktického testu slouží koeficient reliability. Tento koeficient v praxi nabývá hodnot od 0 (pro případ naprosté nespolehlivosti a nepřesnosti) až po hodnoty blízké 1 (pro případ dokonalé spolehlivosti a přesnosti didaktického testu).

Pro individuální pedagogickou diagnostiku se většinou požaduje koeficient reliability minimálně 0,80. Reliabilita testu je odvislá od kvality testových úloh,

z nichž je vytvořen, a závisí také podstatně na jejich počtu. Obecně platí, že čím více úloh test obsahuje, tím má větší reliabilitu. U testů s malým počtem úloh (např. 10 nebo méně) koeficient reliability zpravidla dosahuje maximálně hodnoty kolem 0,60. Z uvedeného vyplývá, že dobrý didaktický test by měl vždy obsahovat dostatečný počet úloh (za spodní hranici lze ve většině případů považovat zhruba 10 úloh).

Reliabilita didaktického testu je velmi důležitým ukazatelem jeho technické kvality. Čím nižší je reliabilita testu, tím skeptičtěji je nutno výsledky testování posuzovat.

Aby didaktický test byl dostatečně validní, musí mít vysokou reliabilitu. Vysoká reliabilita didaktického testu však ještě není zárukou toho, že test bude validní. Test může spolehlivě a přesně (reliabilně) měřit určité vědomosti nebo dovednosti i tehdy, jestliže měří něco úplně jiného, než měřit má. Například budeme-li u žáků základní školy didaktickým testem z matematiky zkoušet, jak umějí řešit kvadratické rovnice, může být toto testování sice vysoce reliabilní, ale jeho validita bude velmi nízká (test zkouší jiné učivo než to, které mají žáci ovládat). Koeficient reliability je možné vypočítat několika způsoby. Podrobnější informace nalezneme čtenář např. v práci Chráska (1998).

7.2.6 Jaké jsou možnosti využívání didaktických testů v praxi?

Z výsledků didaktického testu by měl učitel získat co nejvíce informací pro hodnocení žáků, ale také (a to zejména) pro optimalizaci svého dalšího pedagogického působení.

Prakticky po každém použití testu by měl následovat tzv. *diagnostický rozbor výsledků žáků*. Při tomto rozboru si učitel všimá především chyb, kterých se žáci dopustili, a hledá jejich pravděpodobné příčiny. Forma diagnostického rozboru výsledků testu závisí na druhu použitého didaktického testu, zejména na druhu použitých testových úloh. Doporučené postupy tohoto rozboru lze nalézt v odborné literatuře (např. Chráska, 1998).

Dosažené výsledky testování třídy je výhodné znázornit také graficky, pomocí sloupkového diagramu zvaného *histogram četností*. Z histogramu četností lze získat představu o rozložení výsledků ve třídě, tzn. informaci o tom, kolik žáků ze třídy v testu dosáhlo velmi dobrých výsledků, výsledků dobrých, průměrných atd. U histogramu četností se na vodorovnou osu nanášejí dosažené výsledky testování (počty bodů) a na svislou osu počty žáků, kteří daného výsledku dosáhli (četnosti). Podrobnosti tohoto postupu lze nalézt opět v literatuře (např. Chráska, 1998).

Klasifikace výsledků testu

Mnoho nejasností je mezi učiteli v otázce převodu bodového hodnocení na klasifikační stupně. Je to vcelku pochopitelné, když uvážíme, že ani v teorii nebyl

tento problém doposud spolehlivě a jednoznačně rozřešen. Uvedeme některé přístupy, které se v této oblasti nejčastěji uplatňují (Chráaska, 1998).

a) Intuitivní přístup ke klasifikaci

Někteří učitelé přistupují k převodu bodových výsledků na klasifikační stupně zcela subjektivně a sami víceméně intuitivně určují, kolik bodů je potřeba na dosažení určité známky. Pokud se jedná o učitele s velkou pedagogickou a odbornou zkušeností, většinou je toto hodnocení vcelku odpovídající.

Někteří odborníci doporučují jako optimální řešení otázky převodu bodového hodnocení na klasifikační stupně normativní přiřazování klasifikačních stupňů, které se provádí na základě posudku skupiny odborníků. Tato technika je ve školní praxi dobře použitelná. Vytvořený didaktický test necháme posoudit co možná největšímu počtu učitelů – odborníků v daném předmětu – s tím, že je požádáme o návrh klasifikační stupnice. Z jednotlivých posudků je možno určit průměr, čímž se eliminují extrémní názory.

b) Klasifikace na základě procenta správných odpovědí

Někdy se při převodu bodových výsledků na klasifikační stupně vychází z *procenta správných odpovědí*, kterého žák v testu dosáhl. Jeden z možných postupů (Sedláčková, 1993) je ten, že pokud žák odpoví v testu:

- 91–100 % úloh správně, klasifikujeme jej stupněm výborně;
- 81–90 % úloh správně, klasifikujeme jej stupněm chvalitebně;
- 71–80 % úloh správně, klasifikujeme jej stupněm dobře;
- 61–70 % úloh správně, klasifikujeme jej stupněm dostatečně;
- 0–60 % úloh správně, klasifikujeme jej stupněm nedostatečně.

Učitelé touto metodou často postupují při hodnocení písemných prací (např. z matematiky). Zadávají např. celkem čtyři rozsáhlejší úlohy a hodnotí takto: všechny vyřešené úlohy – výborně, tři vyřešené úlohy – chvalitebně, dvě vyřešené úlohy – dobře, jedna vyřešená úloha – dostatečně, žádná vyřešená úloha – nedostatečně.

Užívání této metody převodu bodových výsledků na klasifikační stupně je v některých případech *diskutabilní*, protože nepřihlíží ke skutečnosti, že úlohy v testu mohou mít rozdílnou obtížnost. Odpoví-li např. žák 50 % úloh správně v testu s velmi obtížnými úlohami, může to znamenat dobrý výkon, zatímco stejný výsledek v testu velmi snadném může být výkonem velmi slabým. Tento způsob klasifikace může vcelku vyhovovat u testů, které neobsahují extrémně snadné, ani extrémně obtížné úlohy.

Hodnocení výkonu žáka podle procenta správně vyřešených úloh se důsledně používá u testů ověřujících (kriteriálních). Tyto testy jsou konstruovány tak, aby zkoušely jen nezbytné učivo, a výkon žáků se hodnotí většinou pouze dichotomicky, tj. *vyhověl – nevyhověl*. Aby žák v testu vyhověl, musí zodpovědět správně zpravidla 80–90 % všech úloh.

c) Klasifikace na základě normálního rozdělení

Bodové výsledky žáků v testu můžeme klasifikovat také na základě *normálního rozdělení četností*. U tohoto postupu se vychází z předpokladu, že výkony dosažené v testu u dostatečně velké skupiny žáků odpovídají tzv. Gaussově křivce. To znamená, že nejvíce je vždy výkonů průměrných a na obě strany od průměru potom četností symetricky ubývá. Jestliže klasifikujeme ve shodě s touto zákonitostí, potom nejvíce žáků klasifikujeme stupněm 3, poněkud méně žáků stupněm 2 a 4 a nejméně žáků stupněm 1 a 5. Mezi jednotlivými autory není shoda v tom, kolik procent žáků by mělo být jednotlivými stupni klasifikováno. Všechny návrhy se však shodují v tom, že rozdělení jsou symetrická kolem stupně 3 a tento klasifikační stupeň má také největší četnost. Jedno z nejčastěji doporučovaných rozdělení je např.:

výborný (1)	7 % nejlepších žáků,
chvalitebný (2)	24 % žáků,
dobry (3)	38 % žáků,
dostatečný (4)	24 % žáků,
nedostatečný (5)	7 % nejhorších žáků.

Jestliže se učitel rozhodne pro klasifikaci podle normálního rozdělení, může k tomu s výhodou používat tzv. percentilové hodnoty. Podrobnosti uvádí např. práce Chráasky (1998).

Některé standardizované testy vyjadřují výkon žáků přímo pomocí klasifikačních stupňů, jiné ho vyjadřují pomocí různých druhů standardních škál, např. percentilové škály, C-škály, T-škály. Poučení o těchto otázkách nalezne čtenář např. v citované práci.

7.3 Evaluace klimatu třídy

Helena Grecmanová

Školní třída je pro nás reprezentantem malé sociální skupiny. Někdy bývá považována za školu sociálního života, jindy za prostředí, v němž se uspokojují aktuální sociální potřeby žáka, nebo za činitele, který ovlivňuje žáky, ale zároveň také za produkt a výsledek činnosti a dispozic žáků (Hrabal, 1989).

Žáci, z nichž se daná třída skládá, se spolu s učiteli, kteří ve třídě vyučují, podílejí na tvorbě jejího klimatu. G. Litwin a R. Stringer (1968) chápou klima jako filtr, přes který se musí dostat objektivní fenomény, a zdůrazňují, že není ani tak důležité, co se děje, nýbrž jak je dění vnímáno. Také P. Piřha a Z. Helus si představují klima jako celkové rozpoložení (1993).

Klima jednotlivých tříd se liší. V některých třídách se můžeme setkat s pozitivním klimatem, které žáky a učitele sblíží a vytváří podmínky pro úspěšné plnění školních úkolů a povinností, jinde pociťujeme negativní klima, které nepříznivě ovlivňuje dění ve třídě, které může v konečných důsledcích vést až